Montana Comprehensive Assessment System Accommodation Validation Study

Analysis of the 2008 Science CRT

Grades 4, 8, & 10

John Bielinski PhD
Montana State Assessment Conference
Billings, MT
January 21, 2011

The Color Test

B L U E

Reading Test

Thew ord sare n otsp ace dcor rectl y.

Valid Testing For ALL Students Requires

STANDARDIZATION
&
ACCOMMODATION

What is the purpose of accommodations?

- To level the playing field

- To eliminate things that interfere with a child's best performance

- To improve validity for students with disabilities

Definition: *A modification to the test which is intended to neutralize extraneous sources of difficulty without altering the meaning of the score.*

Extraneous difficulty
- results from barriers imposed by standardization
- differentially impact certain student populations

Accommodations
- should boost scores for students who need them
- should not affect test score validity

Accommodation Study Overview

Purpose

- determine how accommodations impact validity
- meet requirements of Dept of Ed Peer Review

Methods

- Internal consistency reliability
- Item validity
- Inter-test correlation
- Differential Item functioning

Accommodation Study Overview

Limitations & Challenges

- Cannot evaluate test score boost

- Cannot isolate the effects of individual accommodations

- Mitigate the effects of confounding variables:
    - demographic
    - test performance

Accommodation Study Overview

| Accommodation	Grade					
	4		8		10	
	N	%	N	%	N	%
Small Group Administration	1142	18	877	22	697	22
Oral Presentation	1042	17	601	15	254	8
Alternative Setting	896	14	390	10	374	12
Change in Personnel Administering	847	13	345	9	378	12
Extended Time (no longer used)	696	11	546	14	300	10
Session Duration	331	5	304	8	296	9
Reduced Distractors	240	4	78	2	40	1
Individual Administration	224	4	97	2	65	2
Dictation	204	3	66	2	11	0
Change in Administered Time	184	3	131	3	110	4
Teacher Presence (no longer used)	119	2	123	3	153	5
Test Directions Support	115	2	133	3	113	4
Test Directions with Verification	113	2	11	0	26	1
Fron Row Seating	25	0	41	1	24	1
Writing Tools	16	0	26	1	12	0
Other Acc. Standard	4	0	24	1	58	2
All 13 Other Acc	93	1	180	5	211	7
	6291		**3973**		**3122**	

Accommodation Study Overview

Comparison Groups

- **SpEd Acc**: SWD, 504 plan, or LEP getting accommodations

- **SpEd None**: SWD, 504 plan, or LEP no accommodations

- **GED Matched**: non-acc. GED examinees matched to SpEd Acc by gender, ethnic minority status, free-reduced lunch status, and test score

- **GED None**: remaining non-acc. GED examinees

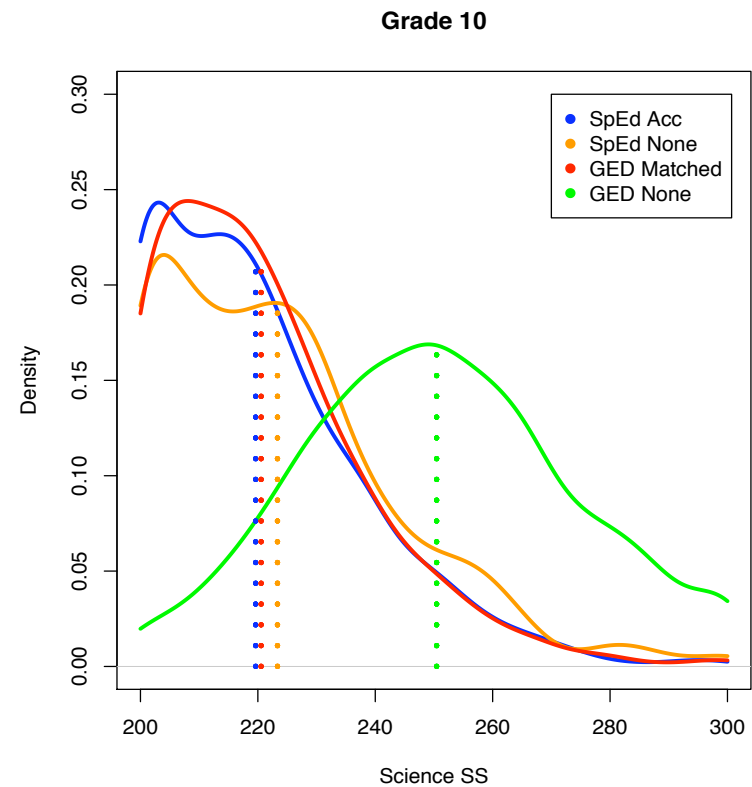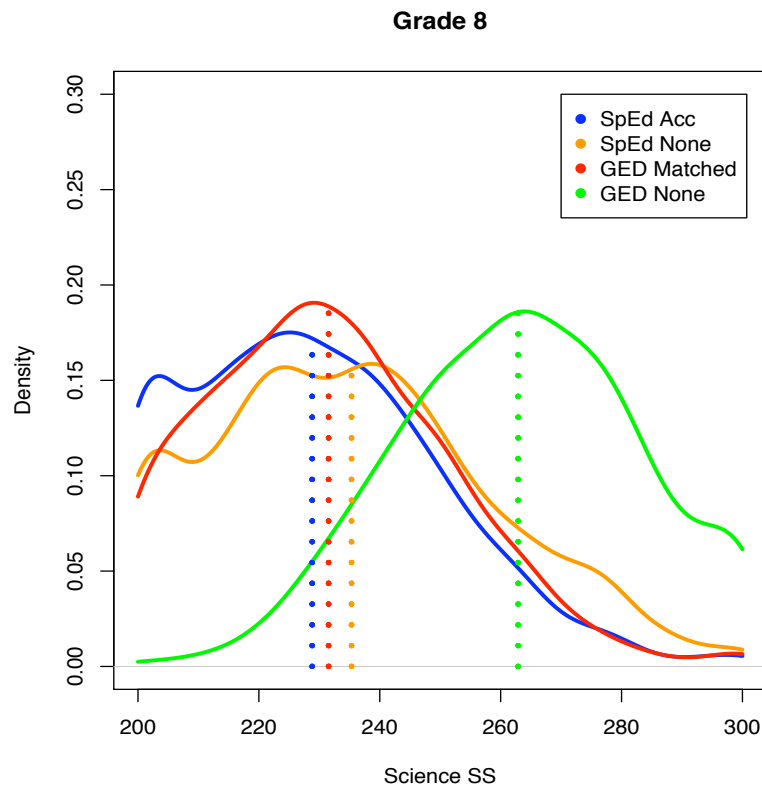| Group	Grade		
	4	8	10
SpEd Acc	851	949	635
SpEd None	704	743	844
GED Matched	823	867	619
GED None	7172	7985	8956
Total	**9550**	**10544**	**11054**

Effectiveness of Matching
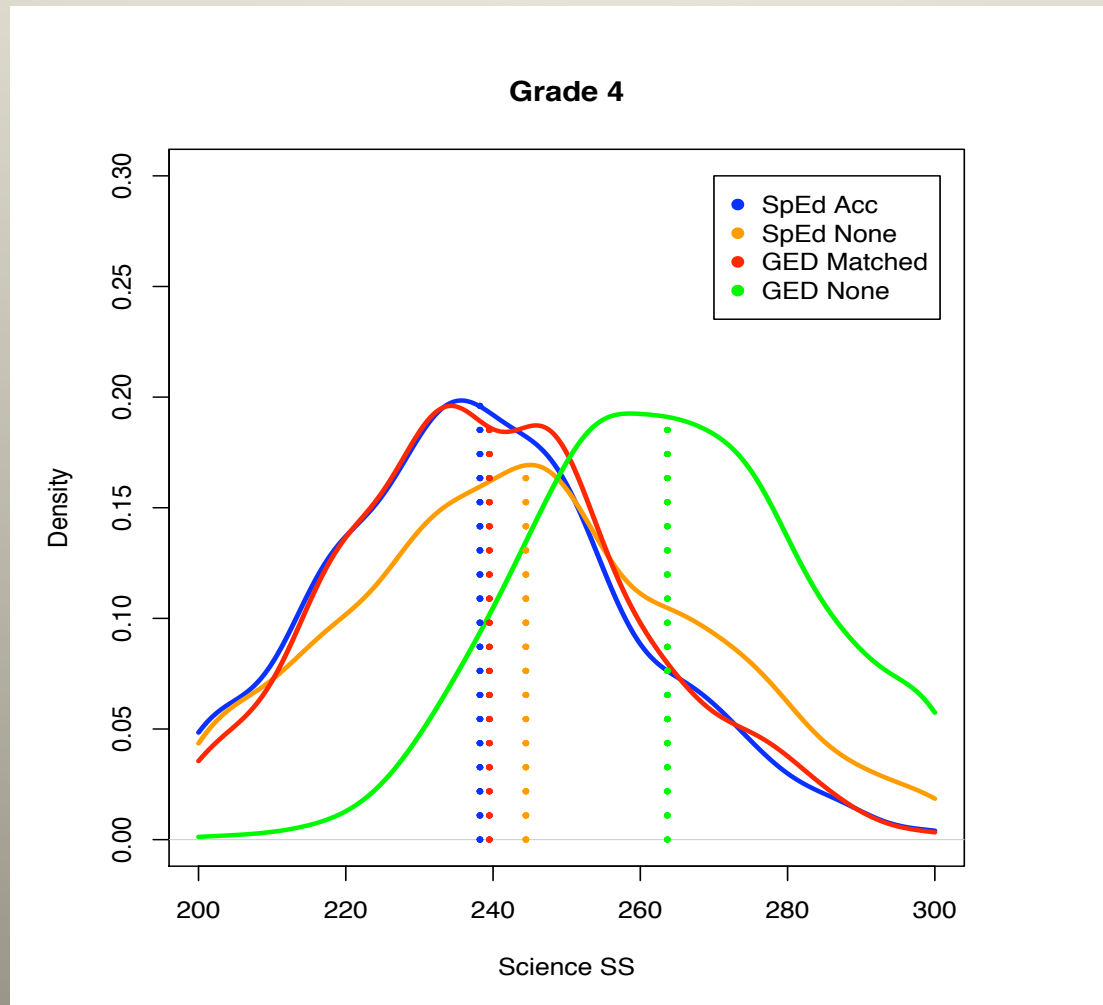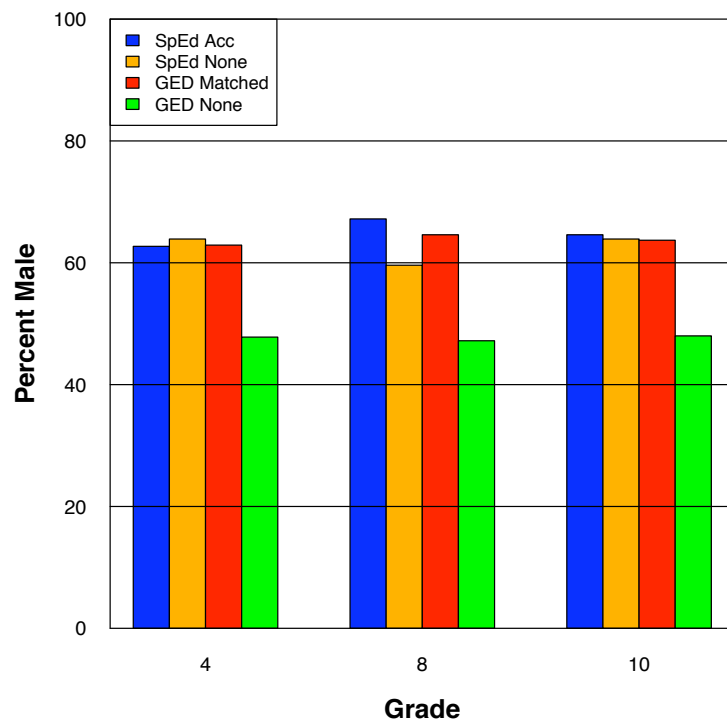
Test Performance

Effectiveness of Matching
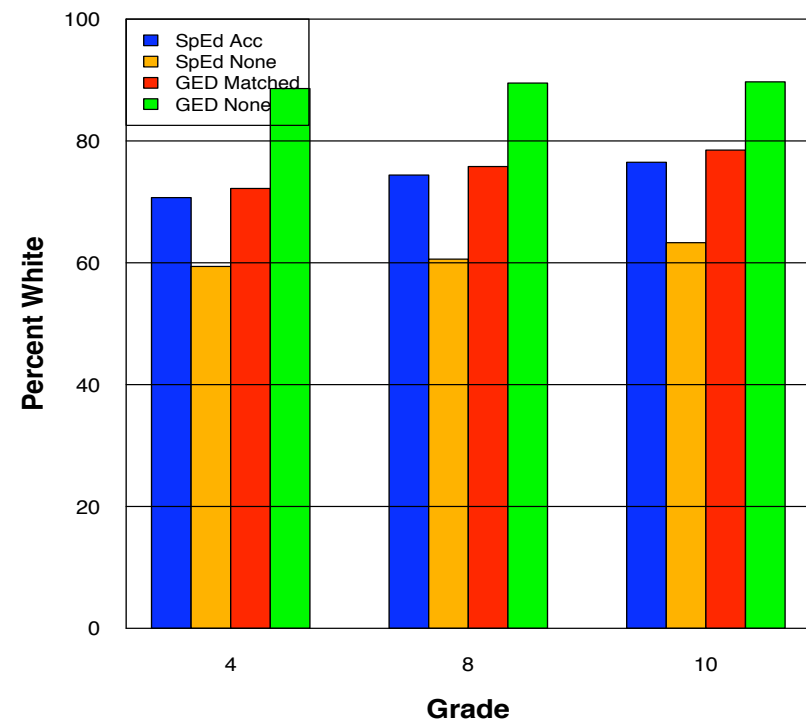
Test Performance cont.

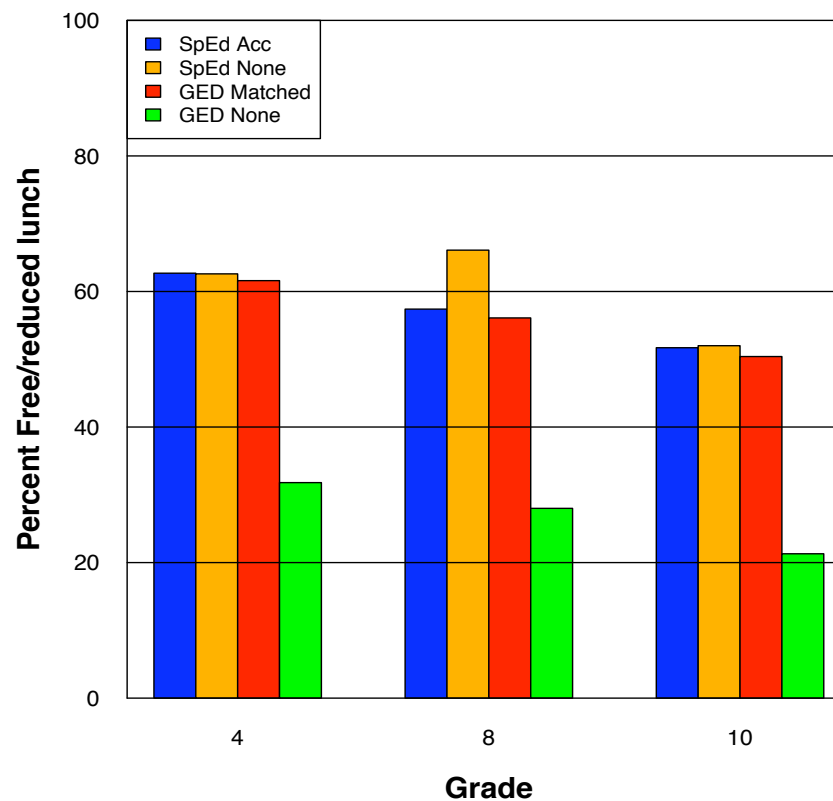# Effectiveness of Matching

Gender

Ethnicity

# Effectiveness of Matching

Free-reduced lunch

Validity Comparisons

Reliability: Cronbach's alpha

| Grade	Group	Cronbach's alpha		
		R_{xx}	Lower Bound	Upper Bound
4	SpEd Acc	0.84	0.83	0.86
	SpEd None	0.88	0.86	0.89
	GED Matched	0.83	0.81	0.85
	GED None	0.79	0.77	0.80
8	SpEd Acc	0.84	0.82	0.85
	SpEd None	0.86	0.85	0.87
	GED Matched	0.82	0.80	0.84
	GED None	0.80	0.78	0.82
10	SpEd Acc	0.81	0.79	0.83
	SpEd None	0.85	0.83	0.86
	GED Matched	0.79	0.77	0.81
	GED None	0.86	0.85	0.88

Validity Comparisons
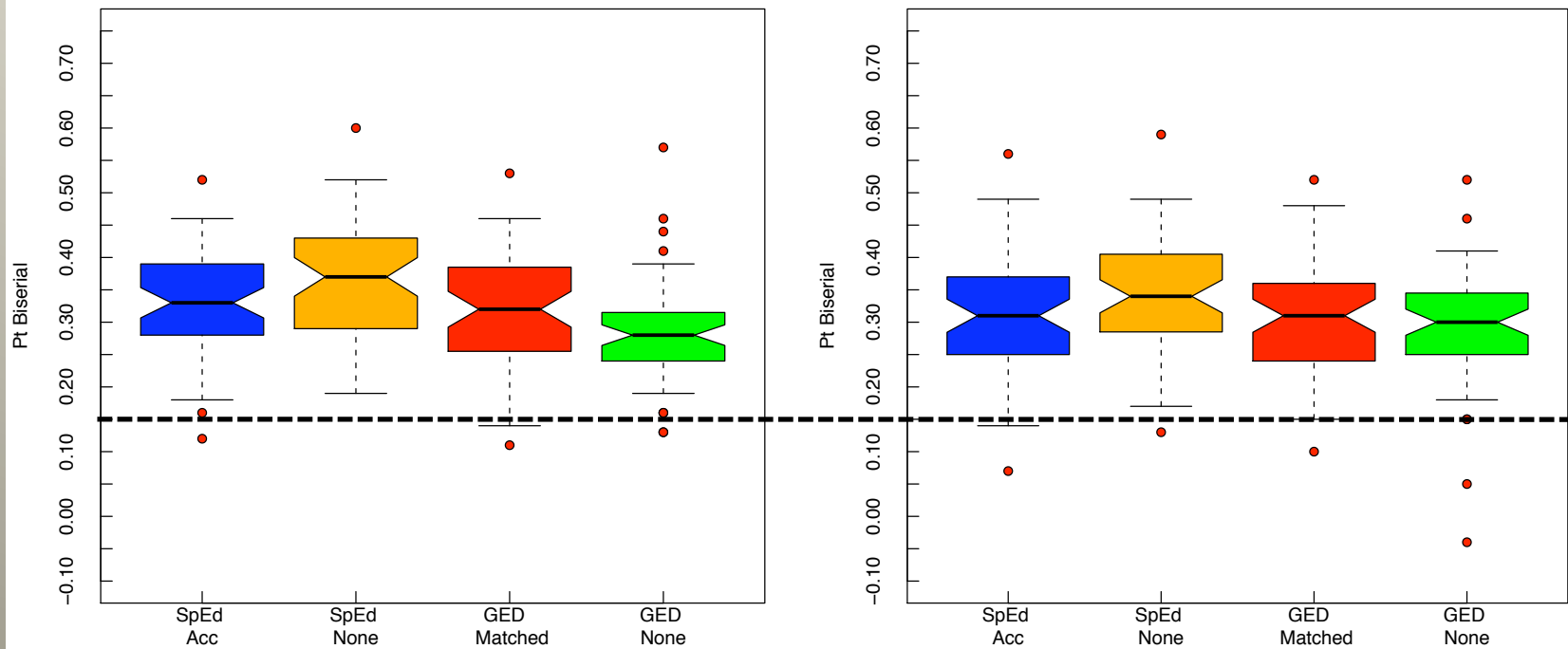
Item Validity: item-total score correlation

Validity Comparisons

Item Validity: item-total score correlation
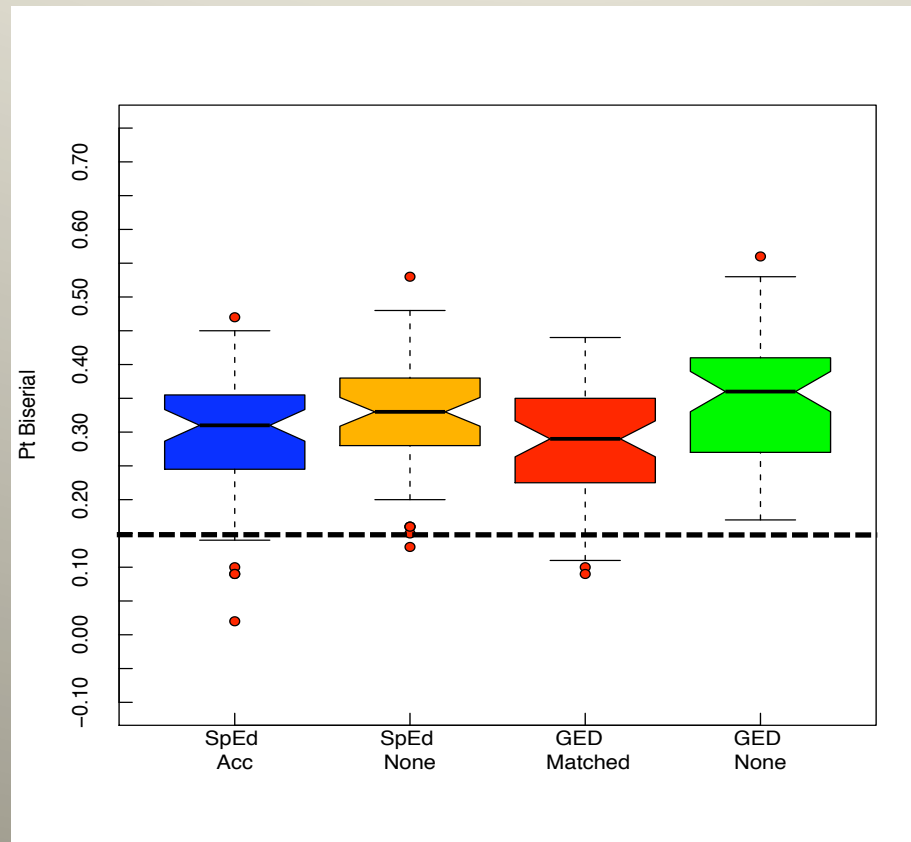
Validity Comparisons

Inter-test Correlation: Reading & Math

Grade	Group	Reading			Math		
		R_{xy}	Lower Bound	Upper Bound	R_{xy}	Lower Bound	Upper Bound
4	SpEd Acc	**0.65**	0.62	0.68	**0.64**	0.61	0.67
	SpEd None	**0.80**	0.77	0.83	**0.77**	0.74	0.80
	GED Matched	**0.73**	0.70	0.76	**0.65**	0.62	0.68
	GED None	**0.69**	0.68	0.70	**0.64**	0.63	0.65
8	SpEd Acc	**0.70**	0.67	0.73	**0.67**	0.64	0.70
	SpEd None	**0.73**	0.70	0.76	**0.73**	0.70	0.76
	GED Matched	**0.67**	0.64	0.70	**0.70**	0.67	0.73
	GED None	**0.63**	0.62	0.64	**0.72**	0.71	0.73
10	SpEd Acc	**0.67**	0.63	0.70	**0.60**	0.56	0.63
	SpEd None	**0.72**	0.69	0.75	**0.70**	0.67	0.73
	GED Matched	**0.57**	0.53	0.60	**0.60**	0.56	0.63
	GED None	**0.67**	0.66	0.68	**0.76**	0.75	0.77

DIF Analysis

Logistic Regression

$$\ln\left[\frac{p_i}{1-p_i}\right] = \beta_0 + \beta_1 tot + \beta_2 group + \beta_3 (tot * group)$$

$\beta_0 + \beta_1 tot$ Model 1 (baseline)

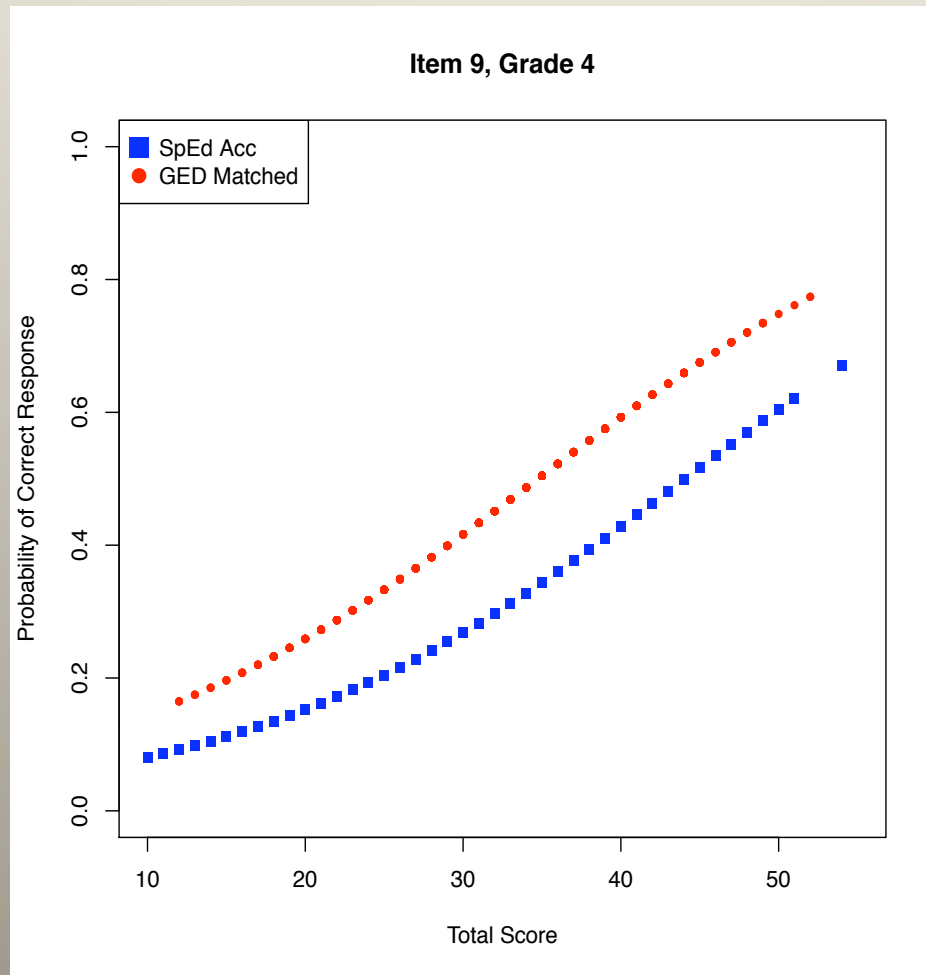$\beta_0 + \beta_1 tot + \beta_2 group$ Model 2 (Uniform DIF)

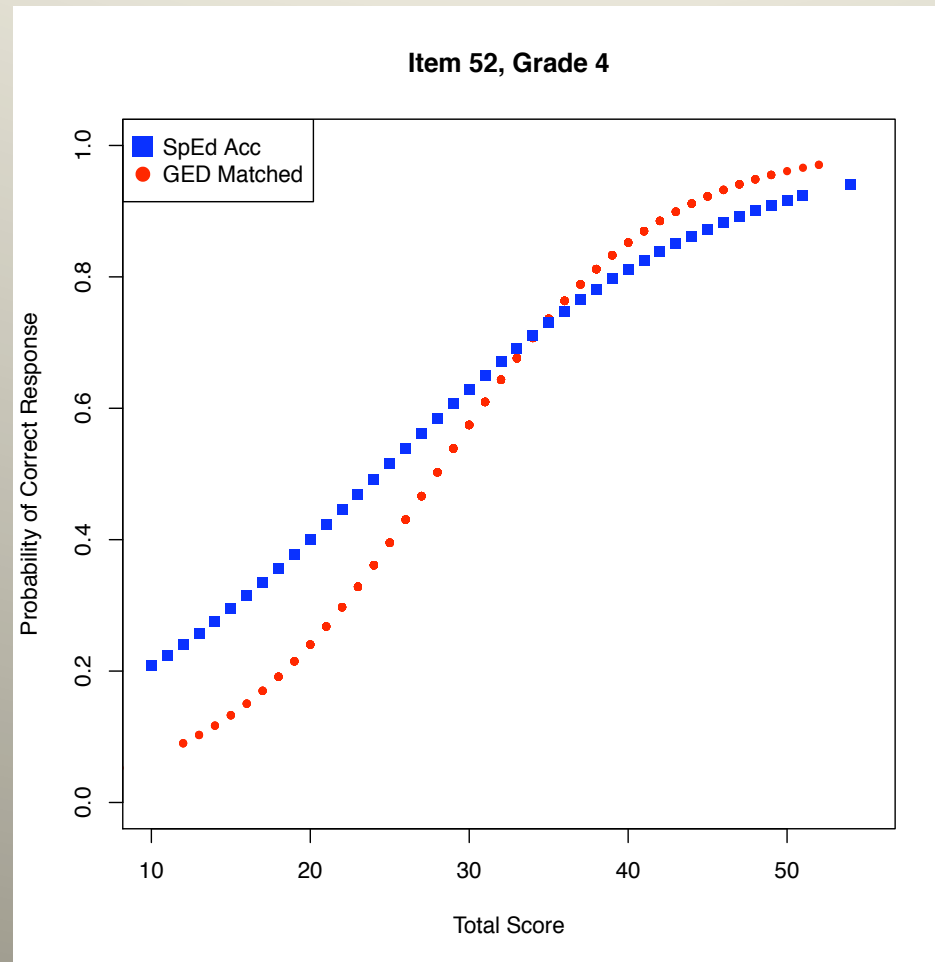$\beta_0 + \beta_1 tot + \beta_2 group + \beta_3 (tot*group)$ Model 3 (Non-uniform DIF)

DIF Analysis

Uniform DIF

DIF Analysis

NonUniform DIF

DIF Results

Uniform & NonUniform

Grade	Focal Group	Uniform DIF				NonUniform		UDIF	NDIF
		FDR	SIB	SIB B	WLS R^2	FDR	WLS R^2		
4	SpEd Acc	1	6	2	4	0	1	**8**	**1**
	SpEd None	1	3	1	1	0	1	**3**	**1**
	GED None	1	0	0	0	14	2	**1**	**14**
8	SpEd Acc	5	8	3	7	0	0	**10**	**0**
	SpEd None	0	2	0	2	0	0	**4**	**0**
	GED None	6	0	0	1	18	2	**6**	**18**
10	SpEd Acc	0	2	1	2	0	2	**2**	**2**
	SpEd None	0	3	2	0	0	1	**3**	**1**
	GED None	8	0	0	0	5	0	**8**	**5**

Conclusions

- SpEd Acc & GED Matched (non-Acc) groups did NOT differ on measures commonly used to validate achievement tests

- Internal consistency reliability, item-validity, and inter-test correlations were as high or higher in the SpEd Acc groups compared to the GED Matched in ALL Grades

- DIF results provide evidence of a modest accommodation affect on item functioning. Overall DIF rates were modestly higher in the SpEd Acc group compared to the SpEd None group

- It is critical that research evaluating the validity of accommodations include multiple demographic and performance matching variables

- Similar results were obtained for the Reading and Math test and for GED students taking the Science test with accommodations